

OBJECT HEIGHT ESTIMATION

by:

Sanjit Dash

(111CS0122)

Under the able guidance of

Prof. Pankaj Kumar Sa

A thesis presented for the degree of

Bachelor of Technology



Department of Computer Science and Engineering
National Institute of Technology, Rourkela

Certificate

This is to certify that the thesis report titled Object Height Estimation in an Image, submitted by Sanjit Dash, of National Institute of Technology, Rourkela for the fulfilment of the requirements for the degree of Bachelor of Technology, is a bonafide record of the work done by him in the department of Computer Science and Engineering under my supervision.

Prof. Pankaj Kumar Sa,
Department of Computer Science and Engineering,
NIT Rourkela.

Acknowledgements

I extend my gratitude to my guide Prof. Pankaj Kumar Sa who guided me throughout the project. His words of advice was very crucial for the completion of the project. His insight into the project made it easier for me to reach the goal. He was a constant source of inspiration who made me do things beyond my limit. Therefore I owe him a lot as I have learnt many things from him which will help in all my future endeavours.

Sanjit Dash
111CS0122
NIT Rourkela

Abstract

This thesis describes methods that mainly deals with the real world height measurement of objects from various images of scenes. The concept of Projective Geometry is used extensively. It forms the basis for accurate estimation algorithms. The thesis is mainly focused on techniques that deals with uncalibrated images; that is the intrinsic properties (focal length and aspect ratio) of the camera or its pose (orientation and position) are not known. There is no need of special calibration device as the geometric characteristics of the scene is extensively used.

Initially the height is estimated by using the pinhole projection formula. But it requires the use of camera intrinsic properties. So we move on to metrology techniques. The geometry of single views is explored and height of the objects in the scene is estimated from the image. To achieve this the properties of planar homographies and planar homologies are extensively exploited. Then properties of vanishing points and lines are also exploited to calculate distances from planes. This also helps in determining the height of the objects in the scene.

The general techniques used in this thesis can be applied to several areas. Examples are presented of commercial, industrial and artistic use.

Keywords: *uncalibrated images, pinhole projection, single view metrology, homography, vanishing points, vanishing line, reference segment.*

Contents

Certificate	i
Acknowledgements	ii
Abstract	iii
1 INTRODUCTION	1
2 LITERATURE REVIEW	2
3 PROPOSED WORK	3
3.1 Pinhole Projection	3
3.2 Simple Scaling	4
3.3 Algorithm to compute the object height present in image (pixels)	4
3.3.1 Assumption	4
3.3.2 Algorithm	5
3.3.3 Matlab functions used	5
3.4 Single View Metrology	5
3.4.1 Measurements on Planes	5
Algorithm to compute height	6
Estimation of the homography matrix H	6
3.4.2 Measurement of distances from planes	6
Vanishing Point and line	6
Algorithm to find vanishing points	7
Algorithm to compute height	7
4 SIMULATION RESULTS	9
4.1 Calculation of height of object in pixels	9
4.2 Height calculation using a reference object	10
4.3 Single View Metrology: Planar measurements	11
4.3.1 Height measurement for a simple image	11
4.3.2 Height measurement for a more complicated image	12
4.3.3 Height measurement for an image consisting of multiple planes	13
4.4 Single View Metrology: Measuring distances from planes	14
4.4.1 Estimating Vanishing Points	14
4.4.2 Estimating Height of object	18
5 CONCLUSION	20

List of Figures

1.1	Perspective distortion in image acquisition process: An image of a wall of Keble College in Oxford, and the four columns have the same real world height, although their heights in the images are clearly not the same due to perspective distortion. . .	1
3.1	Pinhole Projection	3
3.2	Image showing the vanishing points and vanishing line.	7
3.3	Image showing the vanishing point of a line in the scene	7
3.4	Height measurement in a single image. The objective is to calculate the human's height with respect to the reference's height (pillar).	8
3.5	The unknown height ratio Z_x/Z_r can be calculated from image parameters only. . . .	8
4.1	Input image	9
4.2	Output images	9
4.3	Input images	10
4.4	Edge mapped images	10
4.5	Input Image - Four points are chosen on the image plane whose corresponding world co-ordinates are known.	11
4.6	Input Image - Four points are chosen on the image plane whose corresponding world co-ordinates are known.	12
4.7	Input Image	13
4.8	Input Image - Four points are chosen on each image plane whose corresponding world co-ordinates are known.	13
4.9	Input image	14
4.10	Extracted line segments using canny edge detection	15
4.11	Estimated vanishing points	15
4.12	Input Image	16
4.13	Estimated vanishing lines	16
4.14	Input Image	17
4.15	Estimated vanishing lines	17
4.16	Input image of a shed - Height of the man is to be determined with respect to the column (reference object)	18
4.17	Extracted line segments	18
4.18	Points on the man and the column are chosen	19

List of Tables

4.1	Result Table for figure 4.6	13
4.2	Result Table for figure 4.7	14

Chapter 1

INTRODUCTION

The objective of the project is to estimate the real world height of objects from the images of those objects. This is a very important problem because sometimes important measurements are required but it would be too difficult, expensive or time consuming to take them manually. This project can be further extended to estimate depth information from the image. 3D reconstruction is a problem in computer vision where the goal is to construct the scene from the 2D image.

If we know the intrinsic properties of the camera like focal length, pixel size, etc. then it is very easy to calculate the actual height of the object in the image. However, if intrinsic properties are not known then visual metrology techniques can be applied. Metrology literally means the scientific study of measurement. There are two types of metrology techniques (single view metrology and multiple view metrology). Single view metrology deals with measuring the three dimensional geometry of a scene from a single uncalibrated image. Multiple view metrology deals with measuring the three dimensional geometry of a scene from multiple images.



Figure 1.1: Perspective distortion in image acquisition process: An image of a wall of Keble College in Oxford, and the four columns have the same real world height, although their heights in the images are clearly not the same due to perspective distortion.

The most important question is why visual metrology is so hard? The answer to the question is during image acquisition process the 3D space is projected onto a 2D plane due to which some amount of data is lost. In order to reconstruct the scene we need to retrieve those lost information from the photographs. In particular, during the stage of image acquisition perspective distortion occurs. For example, objects which are close to the camera look larger than the objects which are far away. Effects of perspective distortion on images of real world objects are shown in figure 1.1.

Chapter 2

LITERATURE REVIEW

In the past few years Computer Vision is mainly focused on visual metrology and 3D reconstruction of scenes from images. Much exertion was required for seeking after such objectives. In general for a complete 3D reconstruction, it is very difficult for a single image to provide enough information. However from the knowledge of some geometrical information like the relative position of points, lines and planes in the scene some metric quantities can be computed. But in order to achieve that, in general, the camera's intrinsic parameters should be known. The intrinsic parameters are: principal point, skew, focal length, aspect ratio. In the case that the camera's intrinsic properties are not known, a number of visual metrology techniques have been created in order to calculate them. This method is known as camera calibration. Without calibrating the camera it is also possible to take measurements. The work of Criminisi has been the most popular in this area. He has developed methods which helps in calculating measurements from single and multiple views.

Chapter 3

PROPOSED WORK

3.1 Pinhole Projection

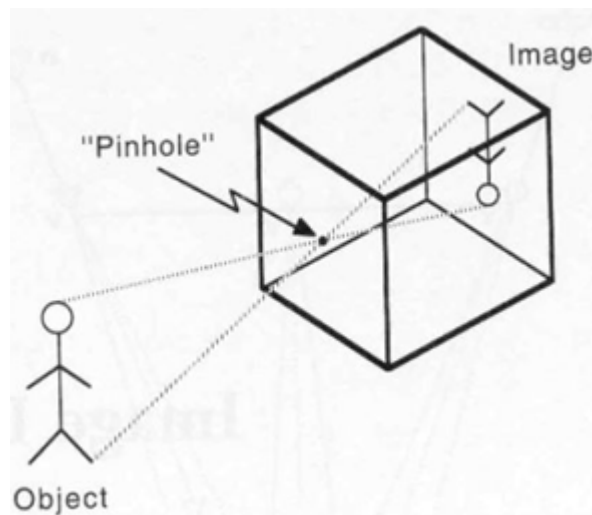


Figure 3.1: Pinhole Projection

The simplest formula to estimate the height of the object is the pinhole projection formula given in equation 3.1.0.1.

$$\frac{x}{f} = \frac{X}{d} \quad (3.1.0.1)$$

where

- x : size of the image.
- f : focal length of the lens.
- X : size of the object.
- d : distance of the object.

If the internal properties of the camera are not known then we can calculate it by using the above formula by taking an object of known height (for example a ruler) and at a known distance.

If d is also unknown then it is difficult for us to tell from the photo if a big object is far away from the camera or a small object very close. This problem can be solved by taking two or more images to calculate the distance.

One approach is to take two photos which lie on the same line with the object. Let the distance to the object from the camera position on the first photo be d_1 and the size of the image be x_1 . Then by applying the pinhole projection formula we get:

$$\frac{x_1}{f} = \frac{x}{d_1}$$

Then the camera is moved s metres directly towards the object, then on the second photo we have size of image x_2 slightly bigger than x_1 :

$$\frac{x_2}{f} = \frac{x}{d_1 - s}$$

which gives us:

$$d_1 = s \times \frac{x_2}{x_2 - x_1}$$

After getting d_1 we can calculate X .

3.2 Simple Scaling

We can also calculate the height of any object with the help of a reference object present in the photo. Suppose for example there is a man and a bottle in a photo. Assuming they are at same distance away and are of the same scale. The actual height of the bottle is known. Let it be h . If 3.5 of this bottle makes the man, then we can get the height of the man (3.2.0.2).

$$Height_{man} = 3.5h \quad (3.2.0.2)$$

This is a simple scaling operation. Without a single point of reference for everything, or if they are all on different scales, a lot more information would be needed.

3.3 Algorithm to compute the object height present in image (pixels)

For the above two methods the image height of the object is absolutely necessary to calculate the real world height of the object. So we need to calculate the height of the object in the image in pixels.

3.3.1 Assumption

The background is taken of a single intensity level. If the background contains a range of intensity levels, then segmentation techniques are required in order to separate the foreground from the background

3.3.2 Algorithm

- Read the image.
- Convert the image to gray image.
- Find the edges of the binary image formed from the above step.
- From the edge map find the maximum and minimum , x and y co-ordinates.
- Then calculate the image height.

After getting the height in pixels, by multiplying the size of each pixel, we can get the height in the same unit as the focal length.

3.3.3 Matlab functions used

edge(img) takes a binary image or an intensity image *img* as its input and returns a binary image which has the same size as *img*, with 1s where edges are detected by the function and 0s elsewhere. By default the *Sobel* method is used by the edge function to detect edges. *ind(I)* returns a vector containing the linear indices of each non-zero element in array *I*.

3.4 Single View Metrology

If neither the camera's intrinsic parameters nor reference object are known, then visual metrology techniques can be applied. Metrology means the scientific study of measurement. There are two types of metrology techniques:

1. Single View Metrology
2. Multiple View Metrology

Single view metrology deals with how the 3D geometry of a scene may be measured from a single image.

3.4.1 Measurements on Planes

An image of a planar surface π is given. Homography is a projective transformation technique that is used to map the points on a 2D image plane to the respective points on the 3D plane. Mapping of co-ordinates in one plane into the corresponding co-ordinates in the other plane as given in equation 3.4.1.1.

$$X = Hx \quad (3.4.1.1)$$

where x is a point in the image plane, X is the respective point on the 3D plane (both expressed in homogeneous coordinates) and the matrix H is a 3×3 matrix having 8 unknowns. The matrix H represents the homography matrix that is used for transformation purpose. Therefore, if we can determine the homography matrix H , we can transform any point in the image plane into its corresponding point on the 3D space and we can get the distances between world points by Euclidean geometry.

Algorithm to compute height

1. First the homography matrix H between the image and world is estimated from a given image of a planar surface.
2. Repeat
 - (a) Two points x_1 and x_2 are selected on the image plane.
 - (b) In order to get the two corresponding points X_1 and X_2 on the world plane, each image point is back projected into the world plane via 3.4.1.1.
 - (c) The distance $d(X_1, X_2)$ is computed between the two world points found by using Euclidean geometry.

Estimation of the homography matrix H

In the case of cameras that are uncalibrated, in order to accurately estimate the homography between the image and the world planes a set of known image to world correspondences needs to be known. Two equations are provided for each image-to-world point correspondence by using the equation 3.4.1.1. They are:

$$h_{11}x + h_{12}y + h_{13} = h_{31}xX + h_{32}yX + h_{33}X$$

$$h_{21}x + h_{22}y + h_{23} = h_{31}xY + h_{32}yY + h_{33}Y$$

For n image-to-world point correspondences we obtain a system of $2n$ equations containing 8 unknowns. So in order to obtain an exact solution 4 correspondences are required that is $n=4$. So in order to get the homography matrix we have to solve 8 equations. Writing the H matrix as a vector

$H = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33})^T$ The homogenous equation 3.4.1.1 becomes $Ah = 0$, with A a $2n \times 9$ matrix.

h is a unit eigenvector of the matrix ATA corresponding to the minimum eigenvalue. From h , homography matrix H is determined.

3.4.2 Measurement of distances from planes

The demerit of the previous method was for every plane we need to know four image to world correspondences to find out the height of any object in each plane. But this method doesn't have that demerit as the height of the object is measured with respect to an object present in the scene. This method involves the use of vanishing points. Here only one reference is taken in the entire image and with respect to the height of the reference object, height of the objects are found.

Vanishing Point and line

In graphical perspective, a vanishing point is a point in the picture plane π that is the intersection of the projections of a set of parallel lines in space onto the picture plane.

Vanishing line is the line joining the two horizontal vanishing points. It is also called the horizon (3.2).

All lines having the same direction share the same vanishing point.

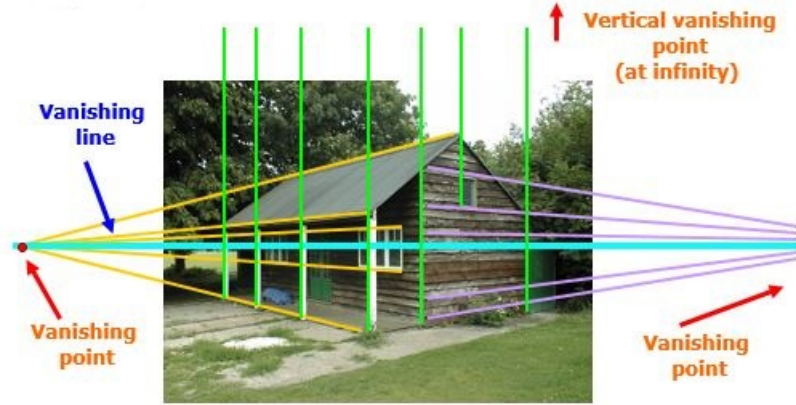


Figure 3.2: Image showing the vanishing points and vanishing line.

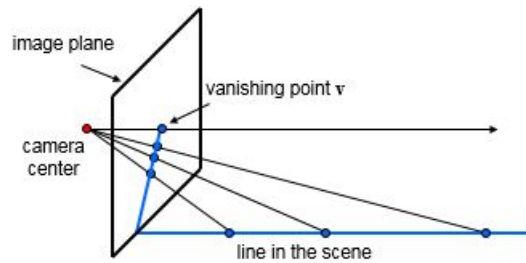


Figure 3.3: Image showing the vanishing point of a line in the scene

Algorithm to find vanishing points

1. First detect the edges using some edge detection algorithm and the set of straight edge segments \mathcal{E} is found by straight line fitting.
2. Repeat
 - (a) 2 segments $s_1, s_2 \in \mathcal{E}$ are selected and they are intersected to give the point p .
 - (b) The support set s_p is the set of all the straight edges in \mathcal{E} intersecting at point p .
3. The point p with the largest support s_p is set as the dominant vanishing point.
4. All edges in s_p are removed from \mathcal{E} and in order to compute the next vanishing point goto 2.

Algorithm to compute height

1. The vertical vanishing point v is calculated.
2. By using the two horizontal vanishing points the vanishing line l , of the reference plane is estimated.
3. The top and bottom points of the reference object (points tr and br respectively) are selected.
4. The metric factor α is calculated by the application of the formula $\alpha = -\frac{\|br \times tr\|}{Zr(l.br)\|v \times tr\|}$
5. Repeat

- (a) Top and bottom points of the object whose height needs to be determined are selected (points t_x and b_x respectively).
- (b) The object's height Z_x is calculated by the application of formula $Z_x = -\frac{\|b_x \times t_x\|}{\alpha(l, b_x) \|v \times t_x\|}$

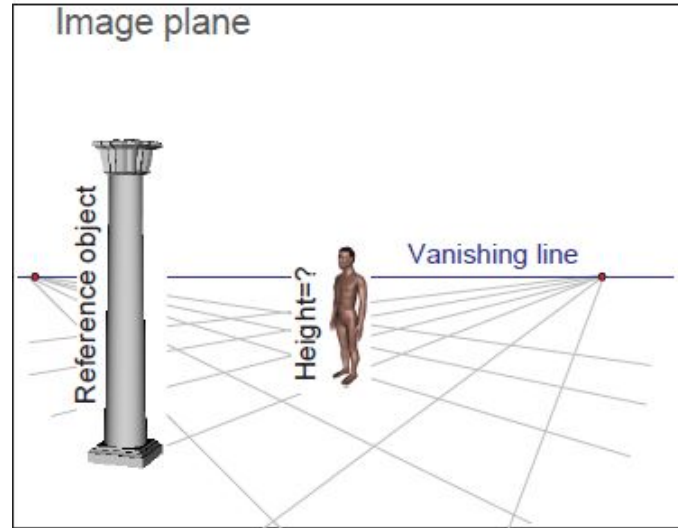


Figure 3.4: Height measurement in a single image. The objective is to calculate the human's height with respect to the reference's height (pillar).

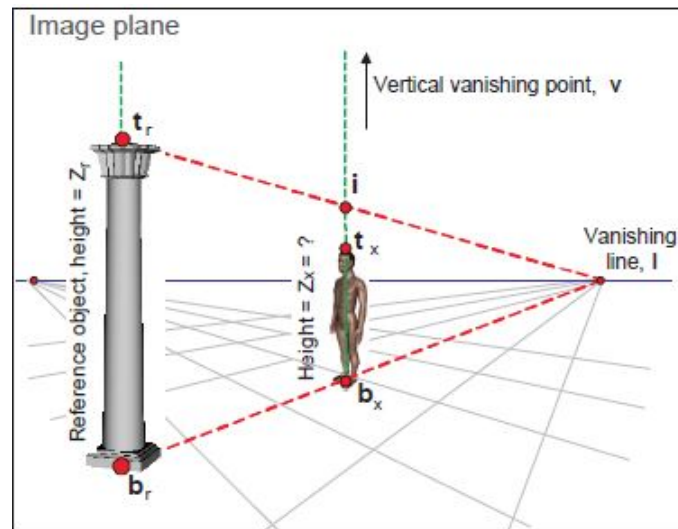


Figure 3.5: The unknown height ratio Z_x/Z_r can be calculated from image parameters only.

So using only the reference's height as a base, the object's height can be found using the above algorithm.

Chapter 4

SIMULATION RESULTS

4.1 Calculation of height of object in pixels

The algorithm mentioned in 3.3.2 was used to get the following result.



Figure 4.1: Input image

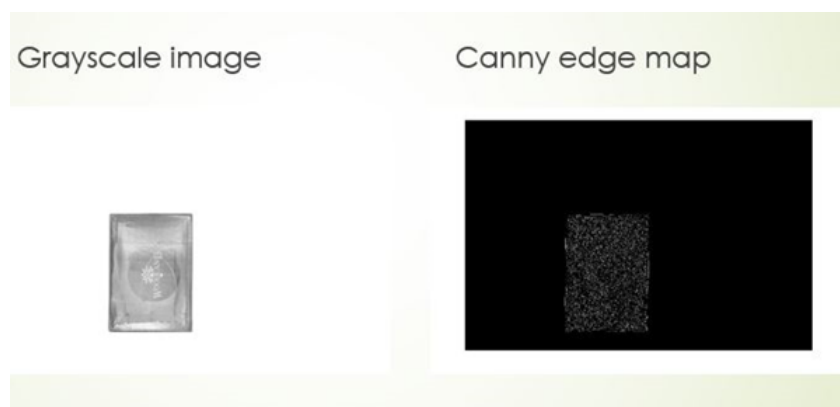


Figure 4.2: Output images

- $Y_{max} = 3006$
- $Y_{max} = 3006$
- Height of the object = 1674 pixels

4.2 Height calculation using a reference object

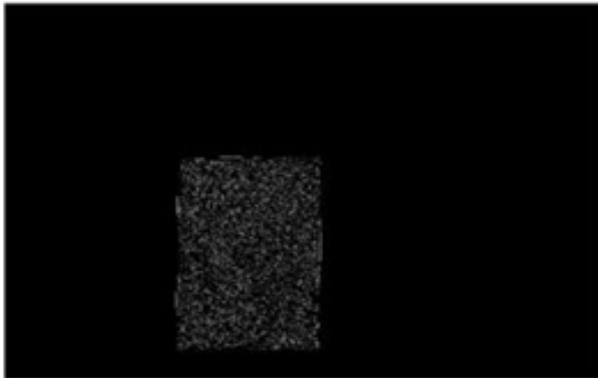


(a) Input without reference

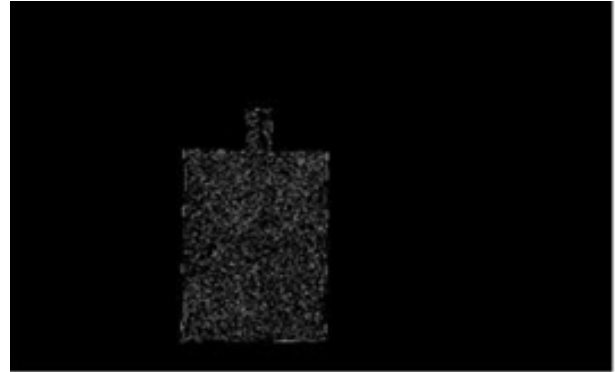


(b) Input with reference. Height of reference is 6.6 cm

Figure 4.3: Input images



(a) Edge map without reference



(b) Edge map with reference

Figure 4.4: Edge mapped images

- $Height_{object} = 1674$
- $Height_{object} + reference = 2027$
- Calculated height = 31.3985 cm
- Original height = 31.5 cm

4.3 Single View Metrology: Planar measurements

Algorithm described in 3.4.1 is used to find out the height of the objects.

4.3.1 Height measurement for a simple image



Figure 4.5: Input Image - Four points are chosen on the image plane whose corresponding world co-ordinates are known.

Four image points and their corresponding world points.

$$(x_1, y_1) = (1880, 2200)$$

$$(x_2, y_2) = (2180, 2200)$$

$$(x_3, y_3) = (2180, 2510)$$

$$(x_4, y_4) = (1880, 2510)$$

$$(X_1, Y_1) = (0, 0)$$

$$(X_2, Y_2) = (5.9, 0)$$

$$(X_3, Y_3) = (5.9, 5)$$

$$(X_4, Y_4) = (0, 5)$$

Using the previous four pairs, the matrix H is calculated and found as:

$$H = \begin{bmatrix} 0.0194 & 0 & -27.5647 \\ 0.0000 & 0.0189 & -25.0868 \\ -0.0000 & 0 & 1.0000 \end{bmatrix}$$

Two image points are chosen and corresponding world co-ordinates are found.

$$x1 = (1420,1330)$$

$$x2 = (1420,3000)$$

$$X1=(0,0)$$

$$X2=(0, 31.4956)$$

$$\text{Height} = 31.4956 \text{ cm}$$

$$\text{Original height} = 31.5 \text{ cm}$$

4.3.2 Height measurement for a more complicated image

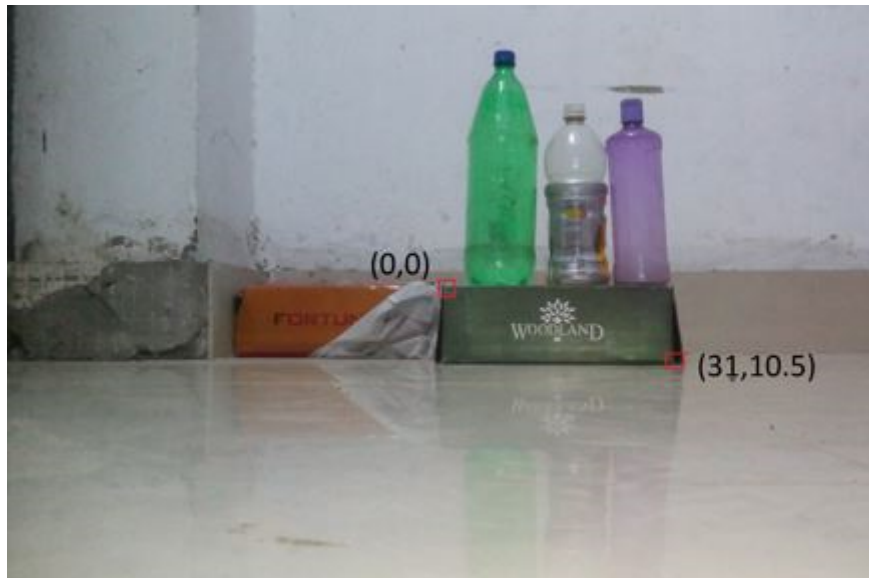


Figure 4.6: Input Image - Four points are chosen on the image plane whose corresponding world co-ordinates are known.

Using the previous two pairs, the matrix H is calculated and found as:

$$H = \begin{bmatrix} 0.0459 & 0 & -59.1797 \\ 0.0000 & 0.0453 & -38.4925 \\ 0.0000 & 0 & 1.0000 \end{bmatrix}$$

Using the homography matrix H the height of the objects are found out.

Sl. No.	Object	Calculated Height(cm)	Actual Height(cm)
1.	Green Bottle	31.1379	33.4
2.	Mazza Bottle	24.0776	26.2
3.	Violet Bottle	24.4397	26.5
4.	Woodland Box	10.3190	10.5

Table 4.1: Result Table for figure 4.6

4.3.3 Height measurement for an image consisting of multiple planes

Here for each plane we will get a separate homography matrix. So for each plane we need four image-to-world correspondences.



Figure 4.7: Input Image

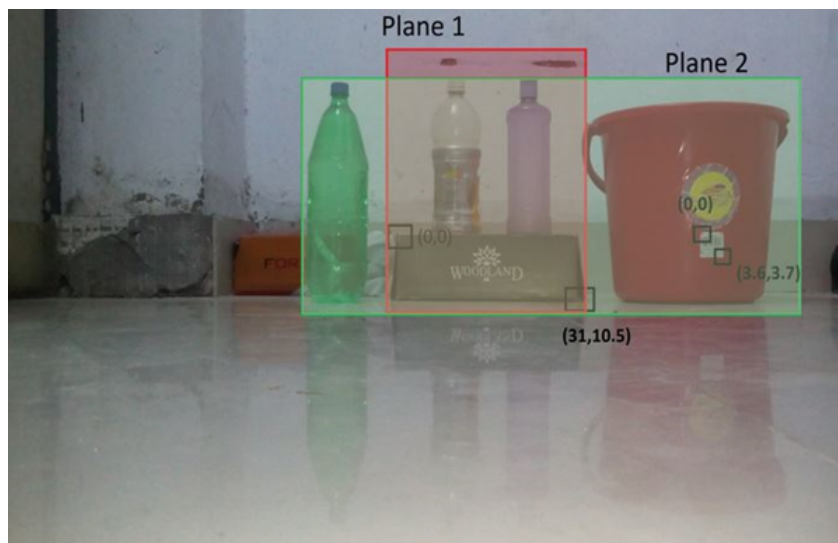


Figure 4.8: Input Image - Four points are chosen on each image plane whose corresponding world co-ordinates are known.

Using the previous two pairs from plane 1, the matrix H_1 is calculated and found as:

$$H_1 = \begin{bmatrix} 0.0500 & 0.0000 & -66.3250 \\ 0.0000 & 0.0486 & -35.3160 \\ 0.0000 & 0 & 1.0000 \end{bmatrix}$$

Using the previous two pairs from plane 2, the matrix H_2 is calculated and found as:

$$H_2 = \begin{bmatrix} 0.0500 & 0.0000 & -66.3250 \\ 0.0000 & 0.0486 & -35.3160 \\ 0.0000 & 0 & 1.0000 \end{bmatrix}$$

Using the homography matrices H_1 and H_2 the height of the objects in their respective planes are found out.

Sl. No.	Object	Calculated Height(cm)	Actual Height(cm)
1.	Green Bottle	32.5600	33.4
2.	Mazza Bottle	23.9167	26.2
3.	Violet Bottle	24.1111	26.5
4.	Woodland Box	9.9167	10.5
5.	Bucket	29.6000	28.8

Table 4.2: Result Table for figure 4.7

4.4 Single View Metrology: Measuring distances from planes

4.4.1 Estimating Vanishing Points



Figure 4.9: Input image



Figure 4.10: Extracted line segments using canny edge detection



Figure 4.11: Estimated vanishing points

$Vp1 = (379, 2333.6)$

$Vp2 = (1204.7, 28)$

$Vp3 = (-200.3, 59)$



Figure 4.12: Input Image



Figure 4.13: Estimated vanishing lines

$$V_{p1} = (248.0, 8470.0)$$

$$V_{p2} = (-27976.8, -118.9)$$

$$V_{p3} = (353.8, 130.6)$$



Figure 4.14: Input Image

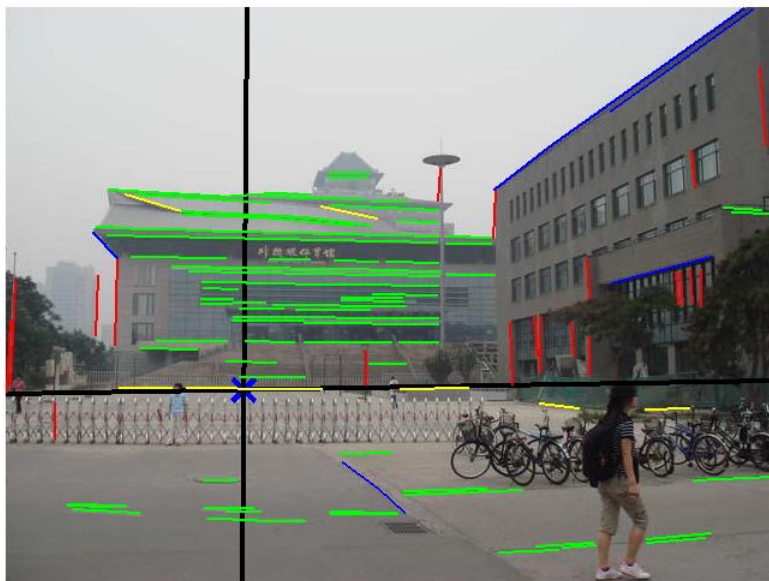


Figure 4.15: Estimated vanishing lines

$$Vp1 = (239.3, -4424.9)$$

$$Vp2 = (3300.8, 265.3)$$

$$Vp3 = (199.0, 319.3)$$

4.4.2 Estimating Height of object



Figure 4.16: Input image of a shed - Height of the man is to be determined with respect to the column (reference object)

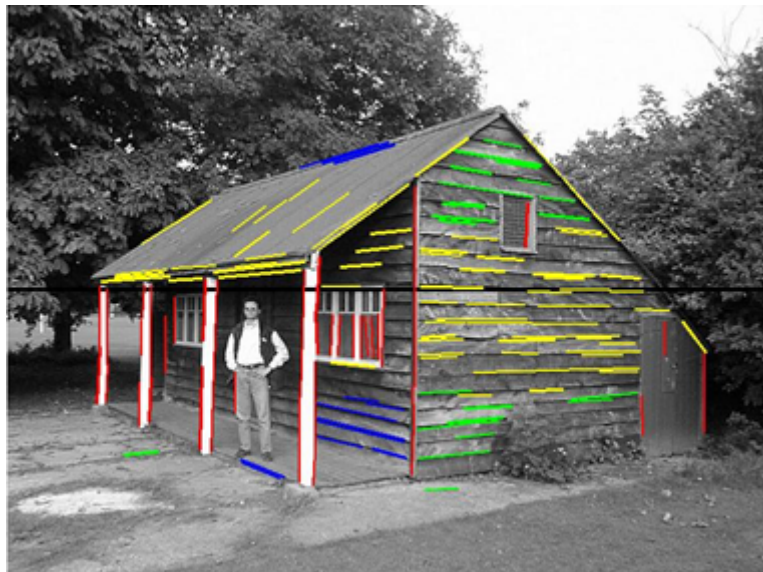


Figure 4.17: Extracted line segments

Vertical vanishing point $V = [0, 100000000, 0]$

Two other vanishing points:

$vp1 = [1867.8, 276, 0]$, $vp2 = [-172.4, 276, 0]$

Therefore vanishing line:

$r = 2040.2$



Figure 4.18: Points on the man and the column are chosen

Actual height of the column, $Z_r = 201$ cm.

Two points on the man is chosen as

$tx = [235.25, 281.75]$, $bx = [236.75, 439.25]$

Two points on the reference is chosen as

$tr = [286.25, 256.25]$, $br = [281.75, 461.75]$

Using the algorithm mentioned in section 3.4.2 we get the following results.

$$\alpha = 1.8135 \times 10^{-13}$$

Calculated Height = 177.7565 cm

Actual Height = 178.8 cm

Chapter 5

CONCLUSION

Object height detection from images is one of the most researched areas in the present image processing research world. Several algorithms and processes have been developed by many research groups to determine the real world height. This study focussed on simple implementation techniques to calculate the real world height of the objects from the image. The concepts of projective geometry have been effectively implemented.

At the beginning of the thesis it has been discussed how to compute the measurements using the pinhole formula. Then the motivation was to compute measurements on planar surfaces. The results of such techniques have been validated with an application suitable for computing measurements from single images of outdoor and indoor scenes.

Metrology algorithms which rely on the camera calibration tend to be affected by such instability arising from variation in temperature or shocks. The algorithms used in the latter half of the report do not make use of internal calibrations, thus are more reliable and robust in their applications. These algorithms can be effectively applied to the existing images as no prerequisites about camera are required.

In this thesis homographies have been used which is a very useful mathematical tool. Homography is a planar projective mapping which is very simple and powerful. During the course of the project we also studied the method of getting affine measurements from single view of images.

This work can be applied to artistic, scientific and commercial applications. This work can be further extended to multiple view metrology. Instead of taking measurements from a single image we can use multiple images for the same purpose. In that case we will get a more accurate result.

Bibliography

- [1] A. Criminisi, i.Reid and A.Zisserman, Single View Metrology, International Journal of Computer Vision, November 2000, Volume 40, Issue 2, pp 123-148.
- [2] Guanghui Wang, Zhanyi Hu, Fuchao Wu, Hung-Tat Tsui,Single view metrology from scene constraints, Image and Vision Computing, volume 23, issue 9, 1 September 2005, pages 831-840.
- [3] A. Criminisi, Single-view metrology: Algorithms and applications, Lecture notes in computer science, 2002, p. 224239.
- [4] A. Criminisi, Accurate visual metrology from single and multiple uncalibrated images, Springer Verlag, 2001.